Expert Systems With Applications 000 (2017) 1–13

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A fuzzy approach for natural noise management in group recommender systems

Jorge Castro [a,b,*], Raciel Yera [c], Luis Martínez [d,e]

[a] Department of Computer Science and Artificial Intelligence, University of Granada, Calle Periodista Daniel Saucedo Aranda s/n, Granada 18071, Spain
[b] School of Software, University of Technology Sydney, PO Box 123, Broadway, Ultimo, NSW 2007, Australia
[c] University of Ciego de Ávila, Carretera a Morón Km. 9 1/2, Ciego de Ávila, Cuba
[d] Computer Science Department, University of Jaén, Campus Las Lagunillas s/n, Jaén 23008, Spain
[e] School of Management, Wuhan University of Technology, Wuhan 430070, China

ARTICLE INFO

Article history:
Received 4 May 2017
Revised 27 October 2017
Accepted 31 October 2017
Available online xxx

Keywords:
Natural noise
Group recommender systems
Collaborative filtering
Fuzzy logic
Computing with words

ABSTRACT

Information filtering is a key task in scenarios with information overload. Group Recommender Systems (GRSs) filter content regarding groups of users preferences and needs. Both the recommendation method and the available data influence recommendation quality. Most researchers improved group recommendations through the proposal of new algorithms. However, it has been pointed out that the ratings are not always right because users can introduce noise due to factors such as context of rating or user's errors. This introduction of errors without malicious intentions is named natural noise, and it biases the recommendation. Researchers explored natural noise management in individual recommendation, but few explored it in GRSs. The latter ones apply crisp techniques, which results in a rigid management. In this work, we propose Natural Noise Management for Groups based on Fuzzy Tools (NNMG-FT). NNMG-FT flexibilises the detection and correction of the natural noise to perform a better removal of natural noise influence in the recommendation, hence, the recommendations of a latter GRS are then improved.

1. Introduction

The Web allows people accessing to a huge amount of information. However, the users skills to cope with all the available information are limited, which leads to select suboptimal alternatives. This problem is known as information overload. Recommender Systems (RSs) are tools to help individuals to overcome such information overload problem personalizing access to information (Adomavicius & Tuzhilin, 2005; Ekstrand, Riedl, & Konstan, 2011). However, some items tend to be consumed by groups of users, such as tourist attractions (Garcia, Pajares, Sebastia, & Onaindia, 2012) or television programmes (Said, Berkovsky, & De Luca, 2011). With this purpose in mind, Group Recommender Systems (GRSs) (Masthoff, 2015) help groups of users to find suitable items according to their preferences and needs.

Several techniques have been used to improve individual recommendation, such as neighborhood-based collaborative filtering (Sarwar, Karypis, Konstan, & Riedl, 2001), matrix factorisation (Koren, Bell, & Volinsky, 2009), or approaches that consider temporal dynamics (Koren, 2010; Rafailidis, Kefalas, & Manolopoulos, 2017). In the case of group recommendation, there are approaches to aggregate individual information (Masthoff, 2015), to consider consensus among members (Castro, Quesada, Palomares, & Martínez, 2015), or matrix factorisation models for groups (Ortega, Hernando, Bobadilla, & Kang, 2016).

A decade ago, it was pointed out that explicitly stated user preferences may not be error free (O'Mahony, Hurley, & Silvestre, 2006). More recently, other recent works (Bellogín, Said, & de Vries, 2014; Centeno, Hermoso, & Fasli, 2015; Guo & Dunson, 2015; Zhang, Zhao, & Lui, 2017) have also pointed out that a person's ratings are noisy, inconsistent, and biased. Li, Chen, Zhu, and Zhang (2013) determined that too many noisy ratings can distort users' preference profiles, which result in *unlike-minded* neighbors that imply a quality loss in recommendations. Kluver, Nguyen, Ekstrand, Sen, and Riedl (2012) have also suggested that user ratings are imperfect and noisy, and such noise limits the predictive power of any RS.

Therefore, in addition to improving recommendations through new recommendation approaches, researchers should also focus on improving the quality of the rating database (Amatriain, Pujol, Tintarev, & Oliver, 2009c). In RSs, there are two kinds of noise in the database (O'Mahony et al., 2006): (i) *malicious noise*, that

* Corresponding author at: Department of Computer Science and Artificial Intelligence, University of Granada, Calle Periodista Daniel Saucedo Aranda s/n, Granada 18071, Spain.
E-mail addresses: jcastro@decsai.ugr.es (J. Castro), ryera@unica.cu (R. Yera), martin@ujaen.es (L. Martínez).

https://doi.org/10.1016/j.eswa.2017.10.060
0957-4174/

consists of erroneous data deliberately inserted in the system to influence recommendations, and (ii) *natural noise*, that appears when users unpurposely introduce erroneous data due to human errors or external factors during the rating process. This paper focuses on the latter.

Natural noise biases recommendations, therefore, its management is a key factor to improve them. There are several Natural Noise Management (NNM) approaches for individual RSs databases. While some NNM approaches need additional information (Amatriain, Lathia, Pujol, Kwak, & Oliver, 2009a; Pham & Jung, 2013), others detect and correct the natural noise using information already contained in the database (Yera, Castro, & Martínez, 2016; Yera Toledo, Caballero Mota, & Martínez, 2015).

GRSs also rely on databases with explicit users' preferences (Masthoff, 2015), therefore, they are affected by natural noise. Castro, Yera, and Martínez (2017) propose a NNM approach for GRSs to manage ratings and noise using crisp values. This is the only work focused on NNM in GRSs. However, the crisp management is not either flexible or robust enough to deal with the uncertainty and vagueness of both the ratings and the NNM, which makes it necessary to develop new proposals with this regard.

In order to manage such uncertainty and vagueness in RSs contexts, the use of fuzzy tools has been considered for several years. A recent survey paper (Yera & Martínez, 2017) has shown that some traditional fuzzy tools have been successfully used for a more flexible and accurate information processing in RSs. However, it also shows that there are several research gaps related to the necessity of new fuzzy approaches focused on the use of emergent information sources and concentrated in new research trends in RSs. Specifically, the natural noise management (Martínez, Castro, & Yera, 2016) is one of such research trends. Our purpose is to study the natural noise management in group recommendation with fuzzy tools.

Therefore, in this work we propose Natural Noise Management for Groups based on Fuzzy Tools (NNMG-FT) to improve the rating database removing the natural noise. NNMG-FT applies three steps of management: fuzzy profiling, global noise management and local noise management. Both global natural noise management step and local noise management step are divided into two sub-steps: noise detection and noise correction. Both sub-steps apply fuzzy tools. In the noise detection, fuzzy tools allow to make a flexible classification of the ratings into noisy or not noisy. In the noise correction, this flexible classification is used to correct noisy ratings applying a soft modification of the value regarding its noise degree. The main advantages of NNMG-FT are: flexibility, robustness and consideration of group information in the NNM. A case study was performed to show the validity of NNMG-FT.

In short, the main contributions of this paper consist of:

- Design an improved profiling that manages uncertainty and vagueness of the ratings through the application of fuzzy tools in the profiling of ratings, users, and items.
- Design an adequate representation of the noise management process that improves the flexibility and robustness of the noise detection and noise correction.
- Propose a NNM approach for GRSs that hybridizes several steps of noise detection and correction based on the information level from the viewpoints of both the whole ratings database and the groups ratings.
- Validate the proposal through comparison with previous ones with similar purpose.

The remainder of this paper is structured as follows. First, Section 2 presents the related works for the current research. Section 3 details NNMG-FT, our proposal for NNM in group recommendation. Section 4 shows the case study done to validate NNMG-FT performance. Finally, Section 5 concludes the work.

2. Related works

In this section we revise different concepts about natural noise management in recommender systems, GRSs, and fuzzy sets, that are used in our NNM approach for GRSs.

2.1. Natural noise management

The existence of underlying noise in users' preferences in RSs and its negative effect have been referred for several years. In this way, an influential paper presented by Herlocker, Konstan, Terveen, and Riedl (2004) pointed out that, although an important amount of advanced algorithms were developed for improving RSs accuracy, the mean absolute error tends to yield around a constant magnitude. They speculated then that such algorithms could be reaching some *magic barrier* where natural variability in ratings may prevent researches from getting much more accurate results. The existence of such *magic barrier* has been confirmed by further investigations in the last few years (Bellogín et al., 2014; Said, Jain, Narr, & Plumbaum, 2012), which have been focused on its characterisation and estimation.

Additionally, the underlying noise in users' preferences began to be referred as natural noise. Formally, natural noise term was first coined by O'Mahony et al. (2006) as those inconsistencies introduced in recommender systems databases due to the imperfect users behaviour when they rate the reviewed or purchased products, without a premeditated malicious intention. It is produced by the influence of external factors in the rating process, such as human errors or rating in different contexts. Natural noise influences the quality of user ratings, and researchers have determined that this influence results in poor recommendations (Amatriain, Pujol, & Oliver, 2009b; Amatriain et al., 2009c). Therefore, an adequate Natural Noise Management (NNM) is key to improve recommendations.

Researchers have explored NNM for individual RSs, which is applied as a preprocessing stage done over the ratings database to reduce the impact of noisy information. Some techniques remove noisy information from the rating database, such as O'Mahony et al. (2006), which deletes both malicious and natural noisy ratings, or Li et al. (2013), which eliminates noisy but non malicious users. These works use the information already contained in the ratings database. However, they overlook important information from the dataset.

There are works that rely on additional information to correct natural noisy ratings. Amatriain et al. (2009c) propose the mining and usage of a curated dataset with information provided by experts to reduce noise. Pham and Jung (2013) uses item attributes to build user models and correct ratings not matching the model, which is built using information of other users identified as experts. More recently, Bellogín et al. (2014) use item attributes for measuring user coherence in recommender systems databases, showing that the recommendation performance is improved when less coherent users are discarded. Later, Yu, Lin, and Yao (2016) propose a correction approach for ratings associated to such less coherent users. Additionally, Saia, Boratto, and Carta (2016) have presented an approach for removing incoherent items from a user profile, using semantic information. These approaches need additional information to correct the noisy ratings, which may not be feasible to obtain in certain domains. Recent proposals also focus on the detection and correction of natural noisy ratings using information contained in the original database. Some of these proposals use contradiction-based approaches (Yera Toledo et al., 2015) or fuzzy tools (Yera et al., 2016).

On the other hand, in RSs context there are items that, because of their social features, tend to be consumed by groups, such as tour packages for groups of tourists (Ardissono, Goy, Petrone,

Please cite this article as: J. Castro et al., A fuzzy approach for natural noise management in group recommender systems, Expert Systems With Applications (2017), https://doi.org/10.1016/j.eswa.2017.10.060